[redacted] – [redacted]
Centre for Health Technology Evaluation
National Institute for Health and Clinical Excellence
Level 1A City Tower
Piccadilly Plaza
Manchester
M1 4BD

Thursday, August 11, 2011

Re: Single Technology Appraisal (STA) - Tocilizumab for the treatment of systemic juvenile idiopathic arthritis – Roche Response to Clarification Questions

Dear [redacted]

Thank you for providing the clarification questions from the Evidence Review Group relating to the clinical and cost effectiveness data of tocilizumab for the treatment of systemic juvenile idiopathic arthritis (sJIA), following our STA submission on 5 April 2011.

Please find below Roche's response to the clarification questions raised on the clinical and cost-effectiveness data.

The requested full clinical study reports, as requested below, may be found on the CD which has been posted out to you today.

If you require any further information or clarification, then please do not hesitate to contact us.

Yours sincerely,

[redacted]

Roche Products Limited [redacted] Tel. [redacted] Fax [redacted]
Registered in England No. 100674

Section A: Clarification on effectiveness data

Requests for additional data

A1 **Priority question:** Please provide full clinical study reports for the TENDER study, as well as MRA316JP, MRA317JP and MRA324JP.

As requested, please find these study reports on an accompanying CD (to follow in the post).

A2 **Priority question:** Please describe which proportion of participants in each trial (especially TENDER) were MTX non-responders; and provide separate data for these participants only (including AEs).

With regards to the TENDER trial, patients are included in the study if they have symptoms of active disease. It follows that if patients have tried in the past or are currently administered MTX and continue to have persistent disease then they are inadequate responders. Therefore, the majority of patients in the TENDER trial are children and young people with inadequate response to NSAID(s), systemic corticosteroids and MTX. In essence, the study design, as presented from the clinical study report, is not stratifying patients based on MTX inadequate response (MTX-IR). Nevertheless, the TENDER study included those patients de facto.

As explained in section 6.2.1 of the MS, there is a small group of patients that are MTX naive (N=5, ~5%), for whom we cannot infer their response to MTX. As stated in the submission, 70% of patients in TENDER were on tocilizumab with MTX and had active disease and these patients can be classified as MTX inadequate responders. The remaining 25% of patients who had MTX in the past, we can assume are also MTX inadequate responders.

In conclusion, the proportion of participants in TENDER that are inadequate responders to MTX is ~95%.

The ACR response at week 12 for the 95% of the TENDER trial who are MTX-IR is presented below. As expected, the results are very similar with that for the 100% of patients.

	TENDER (100%)		TENDER (95%)	
	TCZ arm	PBO arm	TCZ arm	PBO arm
ACR 30	90.7%	24.3%	90.7%	24.3%
ACR 50	85.3%	10.8%	84.0%	10.8%
ACR 70	70.7%	8.1%	69.4%	8.1%
ACR 90	37.3%	5.4%	36.0%	5.4%

A similar trend can be expected for all other study outcomes. This also holds for safety outcomes. However in this case, response to treatment is not an appropriate stratification of patients. A stratification based on concomitant medication (combination of tocilizumab+MTX) would seem more appropriate given the

importance of current treatments to patient tolerability.

We have also looked at the other studies mentioned in question A1 (MRA316JP, MRA317JP and MRA324JP), which were excluded from further discussion for the reasons outlined in our submission. These studies do not provide any subanalysis of MTX response/MTX non-response.

A3 **Priority question:** Please provide similar tables as table 23, page 106 MS, for all other relevant outcomes mentioned in the NICE scope (including AEs), with data separately reported for children and young people 2 years and older who used MTX and those who did not.

In the Week 12 TENDER clinical study report the primary efficacy endpoint 'JIA ACR30 response and absence of fever at Week 12' and secondary efficacy endpoints 'JIA ACR30/50/70/90 responses at Week 12', were investigated with respect to subgroups of use/non-use of background MTX treatment at Baseline (see Table 25 of the original submission [redacted]

In addition, the incidence of adverse events (AEs) to Week 12 was assessed with respect to these groupings [redacted]

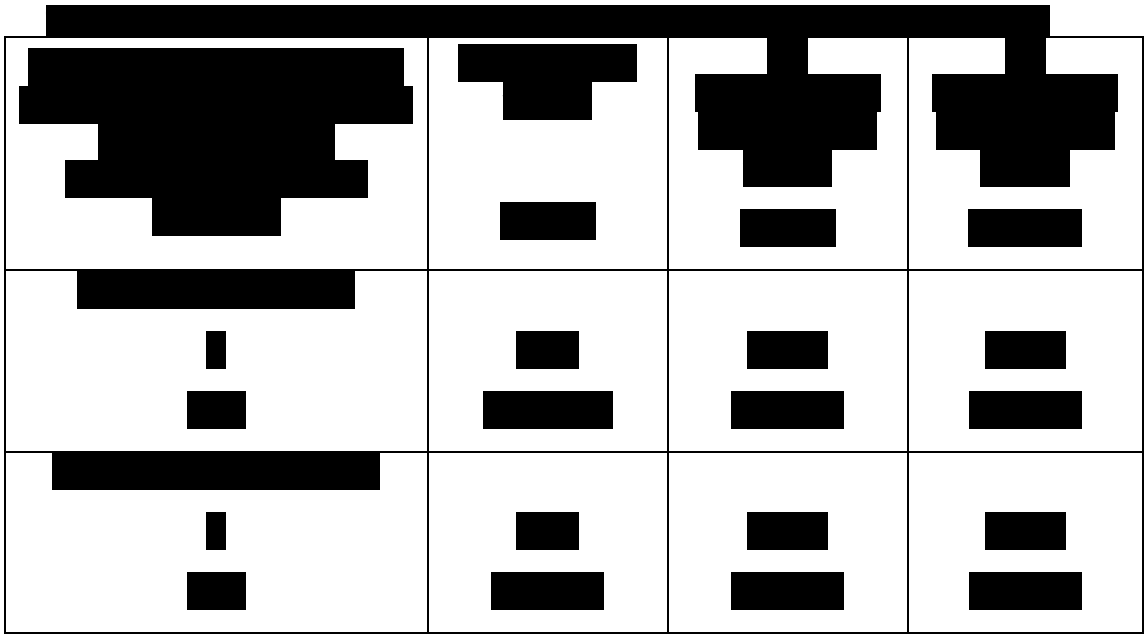
The above described analyses were repeated in a longer-term extension cut, at a timepoint when 50 patients had received TCZ for 1 year in study part II, and are included in a later clinical summary of efficacy and safety documents.

The remaining secondary efficacy endpoints at Week 12, influence on concomitant medications (e.g. corticosteroids) and QoL have not been investigated with respect to subgroups of use/non-use of background MTX treatment at Baseline.

Treated with and without methotrexate

A4 **Priority question:** Looking at the response for placebo without MTX (45%) versus the response for placebo with MTX (15%) (see table 23, page 106 MS), it looks like MTX is harmful in these patients; is this interpretation correct?

Clarification question A4 relates to the following table:



The patient numbers in the placebo arm are quite low and hence it is very difficult to draw conclusions from this even smaller subgroup; it is just that proportionally the placebo patients without MTX appear to have a better response.

We know that there is a lack of evidence to indicate that MTX is superior to placebo in the treatment of sJIA, due to a minimal effect on the systemic features and active arthritis, but based on these low n numbers, we cannot draw any meaningful conclusions that patients taking MTX have a worse response.

Past and present treatments and baseline demographic/disease characteristics must be taken into account to build a picture as to why some placebo patients respond and some do not.

A5 **Priority question:** Please provide a list of patient characteristics and relevant outcomes (including AEs) mentioned in the NICE scope at baseline for children and young people 2 years and older with MTX and without MTX.

In the Week 12 TENDER clinical study report, the primary efficacy endpoint 'JIA ACR30 response and absence of fever at Week 12' and secondary efficacy endpoints 'JIA ACR30/50/70/90 responses at Week 12', were investigated with respect to subgroups of use/non-use of background MTX treatment at Baseline (see

Table 25 of the original submission [redacted]

In addition, the incidence of AEs to Week 12 was assessed with respect to these groupings [redacted]

The above described analyses were repeated in a longer-term extension cut, at a timepoint when 50 patients had received TCZ for 1 year in study part II, and are included in a later clinical summary of efficacy and safety documents.

Hence in terms of a list of patient characteristics and relevant remaining outcomes at Week 12, such as the influence on concomitant medications (e.g. corticosteroids/'steroid sparing') and health-related QoL, these have not been presented in the clinical study report as a separate subanalysis with respect to subgroups of use/non-use of background MTX treatment at Baseline.

A6 **Priority question:** The indirect comparisons described in section 5.7 are for population 2 as described in the NICE scope: "Children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAID(s), systemic corticosteroids and methotrexate." However, the data from the TENDER trial used in the analyses (table 36, page 141 MS) are for the combined population of children (with and without MTX), irrespective of MTX response. Please provide the relevant data (including AEs) from the TENDER trial for children and young people 2 years and older who have not responded adequately to prior NSAID(s), systemic corticosteroids and methotrexate – i.e. not just children and young people 2 years and older who did not use MTX, but children and young people 2 years and older who did not respond adequately to MTX.

The majority of patients in the TENDER trial are children and young people with inadequate response to NSAID(s), systemic corticosteroids and methotrexate. All patients who have already started on MTX prior to entering the study are also MTX inadequate responders. Patients are included in the study if they have symptoms of persistent disease. It follows that if patients were administered MTX and continue to have persistent disease then they are inadequate responders.

As explained in section 6.2.1 of the MS there is a small group of patients that are MTX naive (N=5, ~5%), for whom we cannot infer their response to MTX. In response to your comment, Roche has undertaken a sensitive analysis of the indirect comparison assuming exclusion of all MTX naive patients. The methods are the same as in the MS (Bucher et al. 1997). The results of the indirect comparison sensitivity analysis are presented below.

Comparison	Outcome	Base-case analysis (TENDER)		Sensitivity analysis (excl. MTX naive)	
		RR	95% CI	RR	95% CI
TCZ vs. ANK	ACR30	2.37	1.10, 5.10	2.27	1.06, 4.85
TCZ vs. INF	ACR30	2.87	1.49, 5.55	2.75	1.44, 5.26
	ACR50	5.35	1.91, 14.97	5.04	1.81, 14.04

	ACR70	4.61	1.16, 18.38	4.33	1.09, 17.20

As expected given the small number of patients involved, excluding the MTX-naive patients from the TENDER study has little impact on the results and the conclusions in terms of significance remain the same.

Additional clarifications on clinical effectiveness

A7 Please provide references for the 6 excluded citations that did not report ACR outcomes and the 4 non-English citations, plus any duplicates relating to these (see Figure 11, page 115 MS).

Not ACR outcomes at 12 weeks (6):

Bechtold S, Ripperger P, Dalla Pozza R, Bonfig W, Häfner R, Michels H, Schwarz HP. Growth hormone increases final height in patients with juvenile idiopathic arthritis: data from a randomized controlled study. Journal of Clinical Endocrinology and Metabolism 2007 Aug;92(8):3013-8.

Brunner HI, Lovell DJ, Finck BK, Giannini EH. Preliminary definition of disease flare in juvenile rheumatoid arthritis. Journal of Rheumatology 2002 May;29(5):1058-64.

Ilowite N, Porras O, Reiff A, Rudge S, Punaro M, Martin A, Allen R, Harville T, Sun YN, Bevirt T, Aras G, Appleton B. Anakinra in the treatment of polyarticular-course juvenile rheumatoid arthritis: Safety and preliminary efficacy results of a randomized multicenter study. Clinical Rheumatology 2009 Feb;28(2):129-37

Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ, Stein LD, Gedalia A, Ilowite NT, Wallace CA, Whitmore J, Finck BK. Etanercept in children with polyarticular juvenile rheumatoid arthritis. New England Journal of Medicine, 2000 Mar 16;342(11):763-9.

Sen HN, Levy-Clarke G, Faia LJ, Li Z, Yeh S, Barron KS, Ryan JG, Hammel K, Nussenblatt RB. High-dose Daclizumab for the Treatment of Juvenile Idiopathic Arthritis-Associated Active Anterior Uveitis. American Journal of Ophthalmology, 2009 Nov;148(5):696-703.e1.

Smith JA, Thompson DJ, Whitcup SM, Suhler E, Clarke G, Smith S, Robinson M, Kim J, Barron KS. A randomized, placebo-controlled, double-masked clinical trial of etanercept for the treatment of uveitis associated with juvenile idiopathic arthritis. Arthritis Care and Research, 2005 Feb 15;53(1):18-23.

Not English (4)

Jordan Larraín, Esther; Lucchini Raies, Camila. Effectiveness of a healing cream with human epidermic growth factor in burned AB pediatric patients at COANIQUEM. 2000

Horneff,G.; Hospach,T.; Dannecker,G.; F÷ll,D.; Haas,J.P.; Girschick,H.J.; Huppertz,H.I.; Keitzer,R.; Laws,H.J.; Michels,H.; Minden,K.; Trauzeddel,R. Updated statement by the German Society for Pediatric and Adolescent Rheumatology GKJR on the FDA's report regarding malignancies in anti-TNF-treated patients from Aug 4,

2009. Zeitschrift f³r Rheumatologie, Aug 2010, vol. 69, no. 6, p. 561-7

Martin,Patr□cia; De,Medeiros; Goldenstein,Schainberg. The role of tumor necrosis factor inhibitors in the treatment of juvenile idiopathic arthritis. Revista Brasileira de Reumatologia, , March/April 2006, vol. 46, no. 2, p. 126-133

Rubbert,Roth; Perniok,Andreas. Treatment of patients with rheumatoid arthritis with the interleukin-1 receptor antagonist Anakinra (Kineret«). Zeitschrift fur Rheumatologie, August 2003, vol. 62, no. 4, p. 367-377.

Duplicate (11)

Useful after methotrexate failure in inflammatory rheumatism. Prescrire International, August 2003, vol. 12, no. 66, p. 127-132

Brunner,H.I.; Lovell,D.J.; Finck,B.K.; Giannini,E.H. Preliminary definition of disease flare in juvenile rheumatoid arthritis. The Journal of rheumatology. 2002, 29, 1058-1064

Ilowite,N.; Porras,O.; Reiff,A.; Rudge,S.; Punaro,M.; Martin,A.; Allen,R.; Harville,T.; Sun,Y.N.; Bevirt,T.; Aras,G.; Appleton,B. Anakinra in the treatment of polyarticular-course juvenile rheumatoid arthritis: safety and preliminary efficacy results of a randomized multicenter study. Clinical Rheumatology, 28, 129-137. 2009

Irons,Jung Yoon; Kenny,Dianna Theadora; Chang,Anne B. Singing for children and adults with cystic fibrosis. 2010. Cochrane Database of Systematic Reviews: Reviews 2010.Issue.5 John.Wiley.& Sons., Ltd.Chichester, UK DOI.: 10.1002./14651858.CD008036.pub2

Lahdenne,P.; Vahasalo,P.; Honkanen,V. Infliximab or etanercept in the treatment of children with refractory juvenile idiopathic arthritis: an open label study. Annals of the rheumatic diseases 62, 245-247. 2003

Lovell,D.J.; Giannini,E.H.; Reiff,A.; Cawkwell,G.D.; Silverman,E.D.; Nocton,J.J.; Stein,L.D.; Gedalia,A.; Ilowlte,N.T.; Wallace,C.A.; Whitmore,J.; Finck,B.K. Etanercept in children with polyarticular juvenile rheumatoid arthritis. 2000. New England Journal of Medicine. 342, 763-769.

Lovell,D.J.; Giannini,E.H.; Reiff,A.; Cawkwell,G.D.; Silverman,E.D.; Nocton,J.J.; Stein,L.D.; Gedalia,A.; Ilowite,N.T.; Wallace,C.A.; Whitmore,J.; Finck,B.K. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. 2000 New England Journal of Medicine. 342, 763-769

Lovell,Daniel; Reiff,Andreas; Ilowite,Norman; Wallace,Carol; Chon,Yun; Lin,Shao; Baumgartner,Scott; Giannini,Edward. Safety and efficacy of up to eight years of continuous etanercept therapy in patients with juvenile rheumatoid arthritis. Arthritis and Rheumatism, May 2008, vol. 58, no. 5, p. 1496-1504

Ota,M.O.; Vekemans,J.; Schlegel-Haueter,S.E.; Fielding,K.; Sanneh,M.; Kidd,M.; Newport,M.J.; Aaby,P.; Whittle,H.; Lambert,P.H.; McAdam,K.P.; Siegrist,C.A.; Marchant,A. Influence of Mycobacterium bovis bacillus Calmette-Guérin on antibody and cytokine responses to human neonatal vaccination. Journal of immunology. 168, 919-925. 2002

Ruperto,N.; Lovell,D.J.; Cuttica,R.; Wilkinson,N.; Woo,P.; Espada,G.; Wouters,C.;

Silverman,E.D.; Balogh,Z.; Henrickson,M.; Apaz,M.T.; Baildam,E.; Fasth,A.; Gerloni,V.; Lahdenne,P.; Prieur,A.M.; Ravelli,A.; Saurenmann,R.K.; Gamir,M.L.; Wulffraat,N.; Marodi,L.; Petty,R.E. A randomized, placebo-controlled trial of infliximab plus methotrexate for the treatment of polyarticular-course juvenile rheumatoid arthritis. 2007 Arthritis and Rheumatism. 56, 3096-3106

Smith,J.A.; Thompson,D.J.; Whitcup,S.M.; Suhler,E.; Clarke,G.; Smith,S.; Robinson,M.; Kim,J.; Barron,K.S. A randomized, placebo-controlled, double-masked clinical trial of etanercept for the treatment of uveitis associated with juvenile idiopathic arthritis. 2005, Arthritis and Rheumatism, 53, 18-23

A8 For the 6 excluded citations that did not report ACR outcomes, were any other relevant outcomes (see NICE scope) reported?

Bechtold et al 2007. Growth hormone increases final height in patients with juvenile idiopathic arthritis: Data from a randomized controlled study.
Outcome: Height.

Brunner et al 2002. Preliminary definition of disease flare in juvenile rheumatoid arthritis.
Outcomes: the 6 core response variables (CRV): for JRA were used to derive flare definitions.

Ilowite et al 2009. Anakinra in the treatment of polyarticular-course juvenile rheumatoid arthritis: Safety and preliminary efficacy results of a randomized multicenter study.
Outcomes: Safety: adverse events (serious adverse events and infectious episodes), proportion of patient testing positive for IL-1ra antibodies by immunoassay and for neutralizing capabilities by cellbased bioassay.

Lovell et al 2000. Etanercept in children with polyarticular juvenile rheumatoid arthritis.
Outcomes:
Juvenile rheumatoid arthritis core set criteria: total no. of active joints; no. of joints with limitation of motion and with pain, tenderness, or both; physician's global assessment of disease severity; patient's or parent's global assessment of overall well-being; score on Childhood Health Assessment Questionnaire (did not report the outcomes at 12 weeks, hence is not applicable for the ACR outcomes of our model); erythrocyte sedimentation rate.
Additional assessments: articular severity score; duration of stiffness (min); pain (on a visual-analogue scale); C-reactive protein (mg/dl); no. of swollen joints; no. of joints with limitation of motion

Sen,H et al 2009. High-dose Daclizumab for the Treatment of Juvenile Idiopathic Arthritis-Associated Active Anterior Uveitis.
Outcomes: two-step decrease in ocular inflammation was assessed according to the Standardization of Uveitis Nomenclature criteria.

Smith et al 2005. A randomized, placebo-controlled, double-masked clinical trial of etanercept for the treatment of uveitis associated with juvenile idiopathic arthritis.
Outcomes: ophthalmic efficacy outcomes; inflammation; occurrence of serious infection or sepsis; severe adverse event

Literature searches

A9 For section 9.2.3 'Date span of search' please provide the issue numbers for the searches of CDSR and CENTRAL, as part of the Cochrane Library search. For example, CDSR Issue 1:2011.

CDSR Issue 3:2011
CENTRAL Issue 1:2011

A10 Please can you can you clarify why the comparators were not included in the search strategy of the outcomes but were included in the search for adverse events.

The Final Scope for the submission stated the intervention as: 'tocilizumab with or without methotrexate'. For this reason the initial search strategy for outcomes (found in 9.2) included only tocilizumab as an intervention/comparator, as any trial without tocilizumab would not have directly addressed the Decision Problem.

As stated in the original submission, a more detailed search of all relevant comparators in sJIA (not just focusing on tocilizumab as the intervention) is captured in Section 5.7 (Indirect comparison) as the search strategy did not identify any head-to-head studies which directly compare tocilizumab with the comparators in the 2 populations outlined in the Decision Problem.

Section 5.9 refers to all comparators, which for the submission were methotrexate in population one and TNF inhibitors and anakinra in population two. Therefore all comparators were included in the search, to adequately define any adverse events.

A11 Please provide the search terms for the ACR & EULAR searches. Please provide URLS for the ACR & EULAR conference abstracts searched in section 9.2.4.

The American College of Rheumatology (ACR) and The European League Against Rheumatism (EULAR) congress abstract sites offer free text searches. Various combinations using the following words were tried: tocilizumab and sJIA, JIA, juvenile, idiopathic. The combination which resulted in the most hits was used, and therefore differ between congresses and years.

Congress	URL	Search terms
EULAR all years:	http://www.abstracts2view.com/eular/index.php	"tocilizumab" AND "idiopathic"
ACR 2006	http://acr.confex.com/acr/2006/webprogram/	"tocilizumab" AND "idiopathic"
ACR 2007	http://acr.confex.com/acr/2007/webprogram/	"tocilizumab" AND "idiopathic"
ACR 2008	http://acr.confex.com/acr/2008/webprogram/	"tocilizumab" AND "juvenile"

ACR 2009	http://acr.confex.com/acr/2009/webprogram/start.html	"tocilizumab" AND "idiopathic"
ACR 2010	http://www.abstracts2view.com/acr/index.php?logout=1	"tocilizumab" AND "idiopathic"

A12 Please provide details of the searching undertaken to perform the rapid review described on page 116. Please provide full search strategies, details of databases, database providers, search date and date span.

As discussed in the MS, the rapid review was not performed in a systematic way. This was a supplementary task performed by Roche outside the scope of the submission. The objective was to identify the pivotal study for the comparators; regardless of systemic feature of the population. The methods involve hand searching of published systematic reviews in this area (more details in the MS).

A13 For the date span of searches for indirect and MTC searches in section 9.2.3, please provide the issues numbers for the Cochrane Library search.

The required information from the search performed on the 28th of March 2011 was not recorded. Roche repeated the review of Cochrane Library. Please refer to the accompanying appendix on literature searches for the details of the new search.

A14 Please provide the full search strategy, including issue number and all terms searched, for the Cochrane Library searches for the indirect and MTC searches in section 9.4.4.

Please see above (question A13)

A15 Please provide the issues numbers for the Cochrane Library search for the date span of adverse events searches in section 9.8.3.

CDSR Issue 3:2011
CENTRAL Issue 1:2011

A16 Please provide URLS for the ACR & EULAR conference abstracts searched for the search strategies for the adverse event searches in section 9.8.4.

EULAR all years: http://www.abstracts2view.com/eular/index.php

ACR 2006: http://acr.confex.com/acr/2006/webprogram/

ACR 2007: http://acr.confex.com/acr/2007/webprogram/

ACR 2008: http://acr.confex.com/acr/2008/webprogram/

ACR 2009: http://acr.confex.com/acr/2009/webprogram/start.html

ACR 2010: http://www.abstracts2view.com/acr/index.php

Section B: Clarification on cost-effectiveness data

Model structure

B1 **Priority question:** Please could you redo the cost effectiveness analyses by restructuring the model in health states defined by CHAQ values and use response to treatment to define transition probabilities. The current economic model does not adhere to conventions in Markov modelling. In a Markov model the health states defined should comprise the full range of conditions that are relevant to a patient population, and the states should be mutually exclusive. In the submission, the health states defined reflect a *change* in a patients' condition (change in CHAQ based on ACR response) instead of the absolute condition of the patient. Change in a patients' condition should be included in a Markov model as a health state transition and not as a health state as such. The consequence of using a change in a patients' condition as a health state is that the Markov states are heterogeneous and not mutually exclusive regarding a patients' condition, depending on the disease variation of the cohort at the start of the model. In fact a patient uniquely defined by absolute CHAQ can be situated in any of the ACR-response defined health states. Besides this, the ultimate objective of the model is to report cost-effectiveness in terms of efficiency of interventions expressed in QALYs. As reported, the utility values are not directly linked to the ACR-response (and thus health states) but indirectly via a CHAQ value. Therefore absolute CHAQ is a more accurate parameter to distinguish health states, but absolute ACR score may also be considered. Another way to address this problem would be to use the current model and redo the analysis including a patient level simulation. This will make it possible to define the cost and utility value of a health state depending on the starting CHAQ-value and change in CHAQ in relation to ACR-response.

The economic model uses absolute CHAQ scores as recommended above. The methods of the model adhere to the Markov modelling conventions of health state cohort homogeneity. This is explained in sections 6.2.4, and 6.3.8 of the MS, where Roche presents that ACR is a relative measure and not considered as an absolute measure of health status. Utility values are assigned to the cohort, based on state membership and according to CHAQ score–an absolute score derived after response to treatment.

The labels of ACR 30, 50, 70 and 90 for the health states are only naming conventions, as "response", or "sick" or "well" would be. An alternative label for the health states could be the resulting CHAQ scores (after change from starting CHAQ). However, labelling the health states by ACR response allows numerous changes in sensitivity analysis for parameters

- Starting CHAQ
- CHAQ change from baseline

without revising the methods description each time. This would be confusing as with every sensitivity analysis the names of the health states would differ.

The assumed estimation of the model cohort health status is explained in more detail below. Roche assumes CHAQ reflects the cohort's condition. The cohort enters the model with a starting CHAQ score. Each level of response (ACR30-90) has a corresponding CHAQ-change-from-baseline as observed from the clinical trial. The starting CHAQ is altered after week 12 based on that CHAQ-change-from-baseline.

An example is given below assuming ACR 30 response:

1. Starting CHAQ is: 1.702
2. ACR 30 response is assumed to trigger a CHAQ change from baseline: -0.4318
3. The assumed health status or disease severity of the cohort in the "ACR 30" health state corresponds to a CHAQ of 1.702-0.4318=1.27
4. A utility is assigned based on the QoL formula presented in the MS and according to a CHAQ score of 1.27 (table 77 in MS)

The recommended structural change for estimating the model output by individual simulation would yield very similar results unless variability to starting CHAQ is introduced. Sensitivity analysis has explored the model output given changes to the starting CHAQ value in section 6.7.7 of the MS.

B2 **Priority question:** Based on Fig 13, page 216 MS, it seems that transitions between ACR-response categories is impossible. Please redesign this figure so it reflects all possible health states and all possible transitions between health states

Transitions between ACR-response categories are assumed impossible in the analysis. Roche has evidence that the proportion of "high" responders (ACR 70-90) increases overtime following the first 12 weeks. However, this evidence is available only for tocilizumab with or without MTX. Roche has no evidence of the changes across the ACR categories for any of the other comparators. Therefore, the analysis assumes patients stay to the same health state unless they change treatment line; as reflected in the figure provided.

Treatment strategies

B3 **Priority question:** In the submission MTX is considered not to have impact in terms of treatment response (p 211: it can be assumed that they have inadequate response). Please provide evidence that MTX treatment is inadequate in these patients. Otherwise, this combination treatment should be regarded as a separate strategy, thus in the first population (failure on NSAIDS and corticosteroids) the option tociluzimab/MTX should be included as a separate strategy. Please provide a comprehensive cost-effectiveness analysis comparing MTX alone vs. Tociluzimab alone vs. MTX/Tociluzimab combination treatment and describe the uncertainty of the cost-effectiveness results in cost-effectiveness acceptability curves.

The comment on p211 regarding inadequate response of patients on MTX refers only to patients entering the TENDER trial and having had previously or are still treated with MTX (see response in question A6). That comment is not related to whether there is an impact of MTX when in combination with TCZ or not.

The analysis assumes that the impact of MTX on treatment response is negligible when in combination with TCZ. Clinical experts suggested to Roche that the addition

of MTX in patients with persistent disease rarely has successful results in their functional status and it is considered by clinicians almost equivalent to placebo treatment [Woo P et al. 2000, Personal communication Wright S. 16/03/2011].

Regarding the strategies in the analysis, Roche considers as the base-case tocilizumab according to its licensed indication (+/-MTX). A sensitivity analysis is presented here that presents the cost-effectiveness of a treatment regimen starting with MTX vs. starting with TCZ monotherapy vs. .starting with TCZ+MTX vs. starting with TCZ+/-MTX.

The ACR response (at Week 12) considered for TCZ monotherapy is:

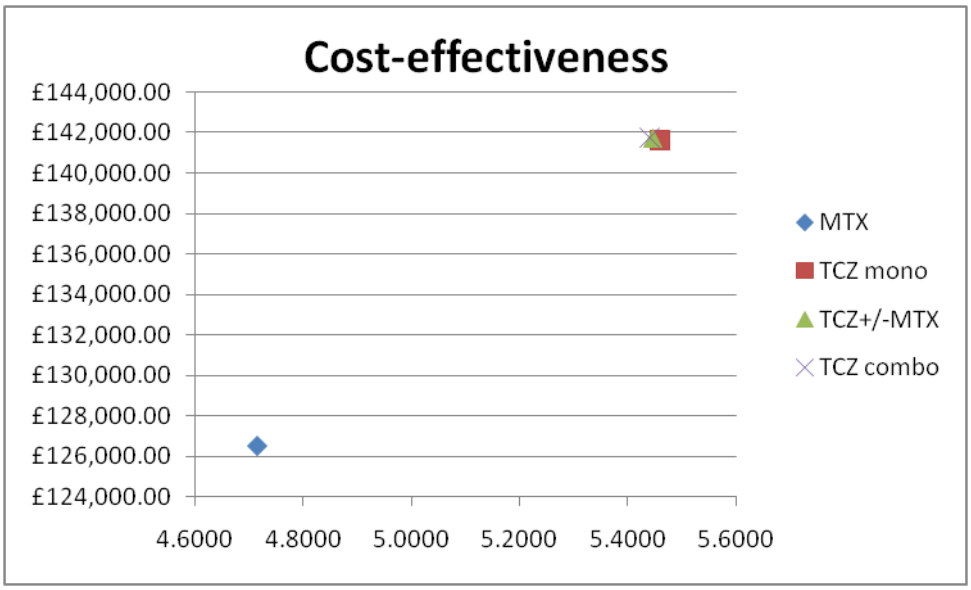
- ACR 30 91.3%
- ACR 50 91.3%
- ACR 70 69.6%
- ACR 90 39.1%

The ACR response (at Week 12) considered for TCZ combination with MTX is:

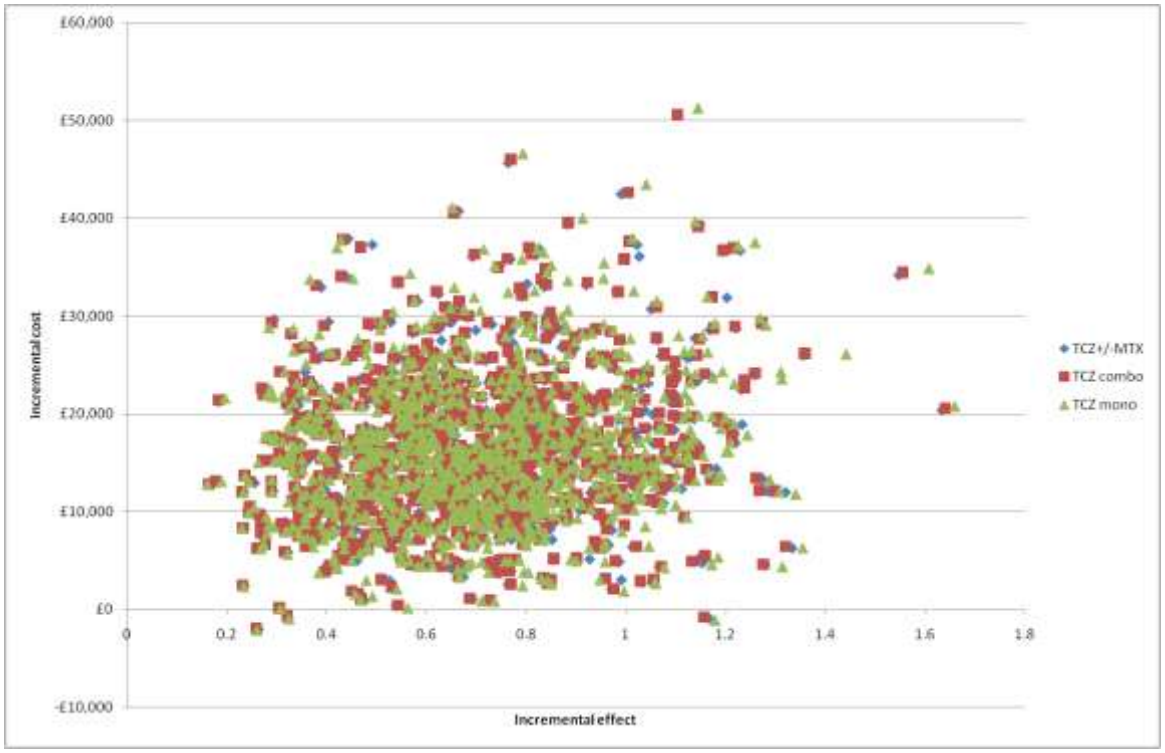
- ACR 30 90.4%
- ACR 50 82.7%
- ACR 70 71.2%
- ACR 90 36.5%

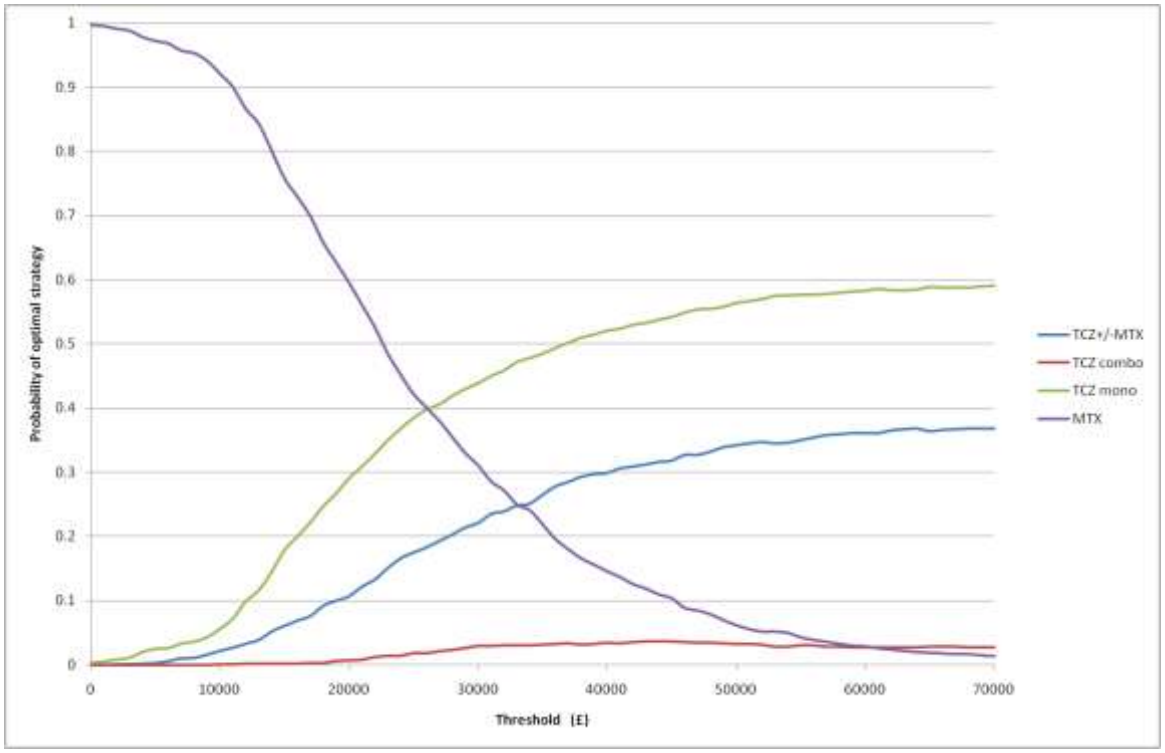
The results of the cost-effectiveness analysis are presented in the table and in the cost-effectiveness plane below. The TCZ strategies have very similar results. The TCZ monotherapy strategy dominates the two combination strategies; per patient it provides more QALYs (0.02 and 0.01 for TCZ combo and TCZ+/-MTX respectively) and costs less (£122 and £90 for TCZ combo ad TCZ+/-MTX respectively). Nevertheless, the differences in results across the TCZ strategies are very small.

Strategy	**MTX**	**TCZ mono**	**TCZ combo**	**TCZ+/-MTX**
Response	3.8270	6.4529	6.4247	6.4341
QALYs	4.7161	5.4594	5.4408	5.4465
Treatment cost	£40,529.21	£89,768.52	£89,439.79	£89,554.10
Health state cost	£85,989.50	£51,858.04	£52,308.89	£52,161.99
Total cost	£126,518.71	£141,626.57	£141,748.68	£141,716.09
Incremental results vs. Standard of care (MTX)				
Incremental cost		£15,107.86	£15,229.97	£15,197.38
Incremental QALYs		0.7434	0.7247	0.7304
ICER		£20,323.46	£21,015.25	£20,806.31



The PSA results presented below include the changes to the Dirichlet distribution as requested from question B8. The first figure presents the incremental cost and effect of the TCZ strategies versus the standard of care (MTX). The second figure reflects the probability of one of the strategies being optimal based on the PSA results and across a range of CE thresholds.





B4 **Priority question:** Please provide observed standard deviations and confidence intervals for the data in tables 74 and 75, pages 213 and 215 of MS.

Please clarify the tables you are referring. Pages 213 and 215 do not include any tables. Table 74 is simply the number of YES out of all the possible 36 criteria of study quality. Regarding table 75; starting age and weight are assumptions (see table 75)–not observed values. The measures of dispersion of the starting CHAQ score are:

Mean 1.702
SD: 0.7930
95%: 1.554, 1.849

B5 **Priority question:** Please provide evidence that treatment efficacy is constant regarding the life expectancy of patients. If not, please provide sensitivity analyses using loss of treatment efficacy.

Roche has no evidence of loss of treatment efficacy over the life expectancy of patients either regarding initial response or withdrawal due to lack of efficacy. In section 6.4.11 of the MS Roche presents evidence of improving CHAQ over time for TCZ patients.

B6 **Priority question:** On page 227 of MS and in figure 14, a constant risk of withdrawal from biologicals is assumed. Please provide full details on other models considered, especially goodness-of-fit statistics. Also, please provide the SEs for the parameter estimates.

In the calculation of the withdrawal risk, a published Kaplan-Meier curve is the only source of information. As ‘raw’ survival data are needed to properly fit a parametric

model, these are estimated from the curve using the following method. Values of the product-limit estimator are read from the KM curves using Grafula version 2.10. The number of events and the number of patients at risk for each time-interval are then derived from the product-limit estimator values using the method developed by Parmar et al. [1998]. For each time-interval, Parmar et al. [1998] start by considering the number of people at risk at the beginning of the interval (starting with the number of people included in the trial). The next step is to estimate the number of people censored during this interval, in order to get the number of people at risk during the time interval. To do this, a model for censoring must be assumed. The model used by Parmar et al. [1998] assumes that the patients are censored at a constant rate during any time interval and thus contribute half a person's worth of risk during the interval. More details on this model and its mathematical aspects are given in the article. Finally, the formula giving the product-limit estimator value is back-transformed to get the number of events for each time interval: as the KM value and the number of people at risk during the interval are known, the number of events is indeed the only unknown in the equation and can therefore be derived.

As a second step, parametric models are fitted to the data and compared using statistical criteria to determine the most appropriate one. As models aren't necessarily nested, the usual Wald/Likelihood ratio tests are not appropriate, so the Akaike Information Criterion (AIC) is used. The AIC is an operational way of trading off the complexity of the model against how well the model fits the data: as shown in the formula above, it rewards goodness of fit but also includes a penalty that is an increasing function of the number of estimated parameters. The preferred model is the one with the lowest AIC value.

In this case, the AIC is calculated as:
$AIC=2(c+p+1)-2\ln(L)$
where

- c is the number of model covariates
- p is the number of model-specific ancillary parameters
- L is the maximized value of the likelihood function for the estimated model.

Both exponential and Weibull parameterizations were fitted to the withdrawal data. The AIC criterion showed that the exponential model was the most appropriate (AIC 177.05 versus 178.74).

Using a standard linear regression approach, the exponential model was then compared to a simple linear model using the R^2 criterion which measures the goodness of fit of the model to the data. Results are shown below:

Linear model
Equation: survival= -0.0948*years + 1
SE around the time parameter: 0.0018
R^2=0.954
The intercept term is set to 1 to ensure that survival is 1 at the time origin and therefore does not have a precision associated with it.

Exponential model
Equation : survival= exp(-0.138*years)
SE around the time parameter: 0.0016
R^2=0.982

Performances of the two models are very similar. Roche identified in the MS (sections 6.3.3 and 6.10.3) that a time-dependent probability (exponential in this case) would be accurate for the withdrawal risk. However, given the evidence on withdrawal risk with MTX and other biologics, it was not possible to differentiate greatly across treatments on this parameter. It was deemed more appropriate from a model parsimony view to use a constant risk (R^2=0.954).

B7 **Priority question:** As previous question but for mortality risk (figure 15): please provide full details on other models considered, especially goodness-of-fit statistics. Also, please provide the SEs for the parameter estimates.

In the calculation of the mortality risk a published Kaplan-Meier curve is the only source of information. However, out of 962 patients and over 14 years of follow-up, there are only 21 events. This is not enough evidence to "recreate" the survival data using the Parmar et al. [1998] method. First, Parmar et al. [1998] relies on an assumption for censoring. With the large number of patients and the small variations in the KM estimate the Parmar et al. [1998] method does not perform well. In turn, this has implications on the estimations of the number of events. Overall the evidence from Hashkes et al. 2010 show that events are few and far between, which does not result in a continuous hazard curve. This does not allow for a fitting of a parametric model. The only models tested here are therefore the linear and exponential ones, using standard linear regression methods. The results are shown below:

OLS FITTING
Linear
survival= -0.0007*years + 1
SE around the time (-0.0007) parameter=0.000015
R^2=0.959

Exponential
survival= exp(-0.0007*years)
SE around the time (-0.0007) parameter=0.000015
R^2=0.958

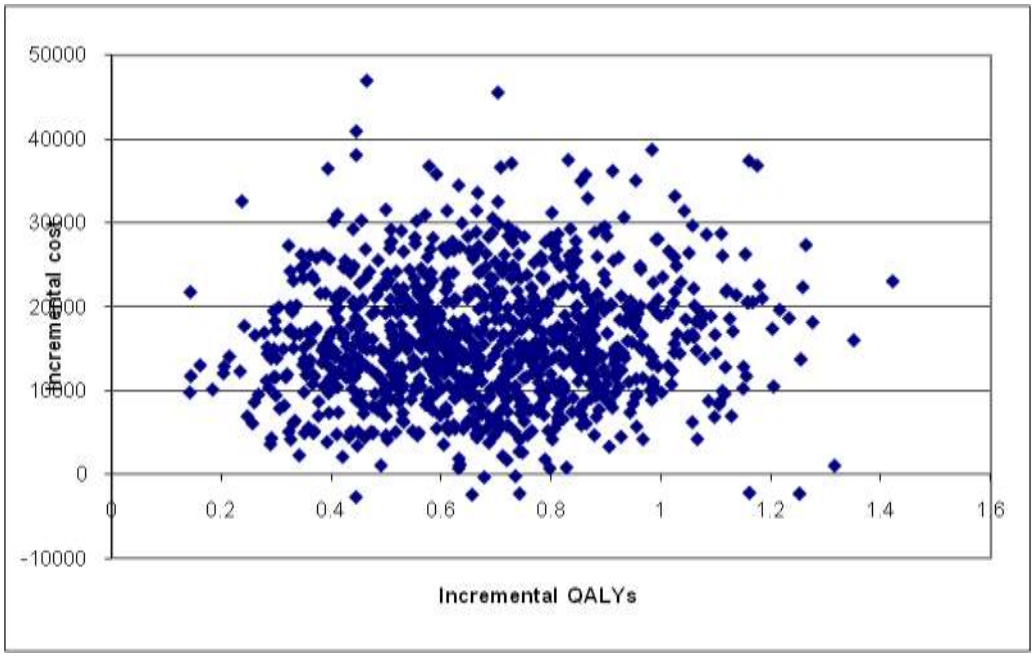
The two models have very similar results. The linear is preferred for the economic analysis.

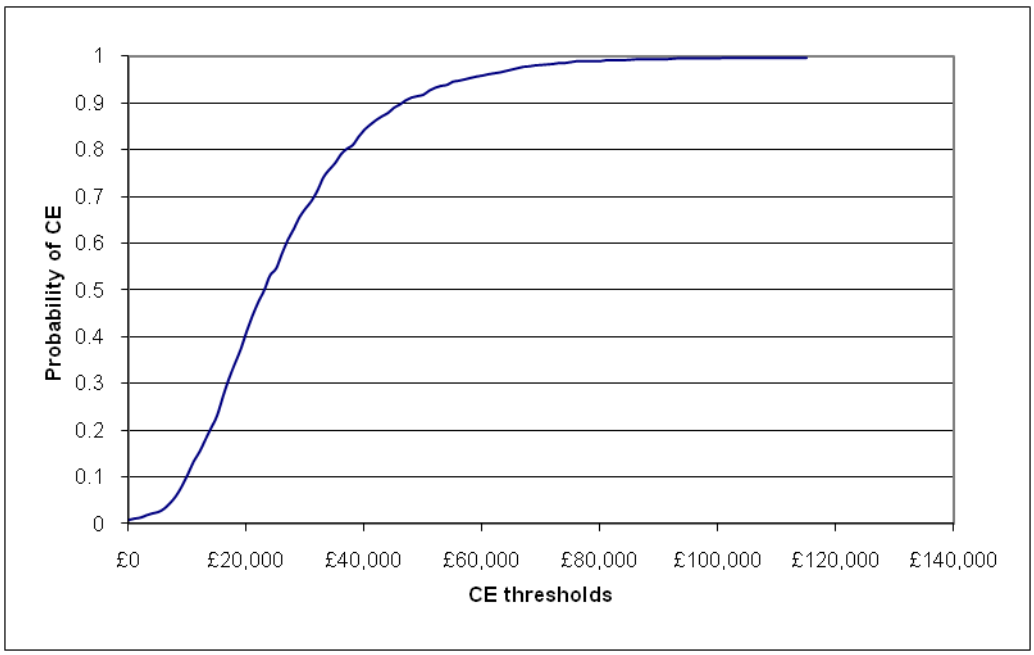
B8 **Priority question:** In table 87, page 236 of MS a summary of variable values is presented, with the distributions used in the PSA. The text in the table for the various transition probabilities is unclear 'assume N=alpha of the one parameter Gamma distribution for each ACR category'. This seems to suggest that the true N is used in deriving random draws from the Dirichlet distribution. However, in the electronic model, N=100 is assumed for all treatments and all transitions. Please change the model input to reflect the true N on which the transition probabilities are based, and provide the PSA outcomes.

The model input was changed to reflect the true N on which the transition probabilities are based. Please note that for the input related to indirect comparison the N of the original trial of the comparator is assumed. The result of the PSA is presented below with scatterplots and CEAC for both comparisons.

TCZ vs. MTX:

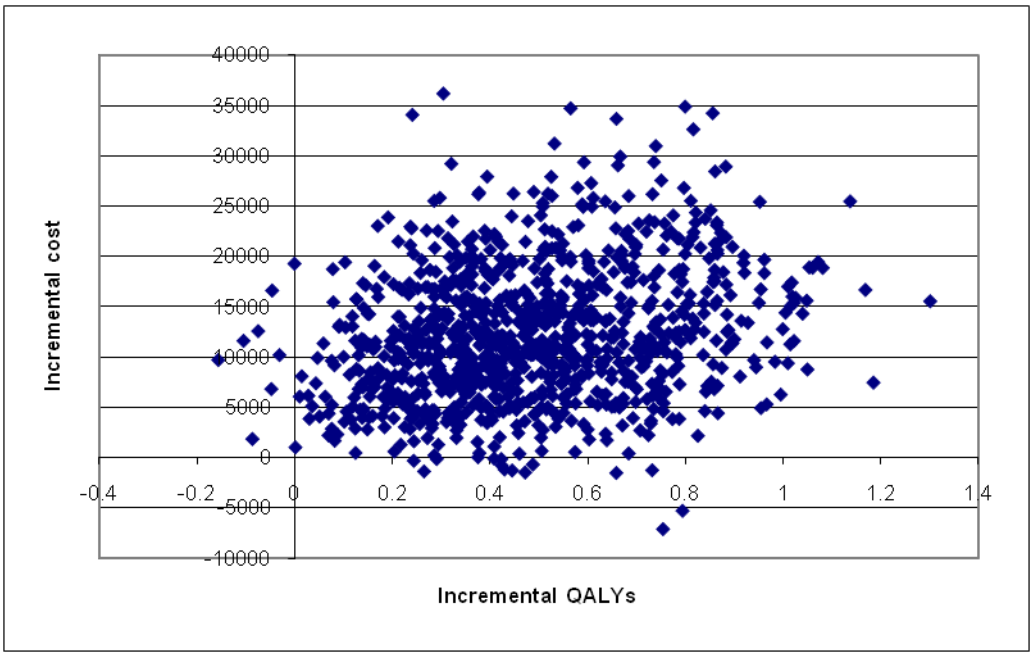
41.1% of samples are below a threshold of £20,000
67.5% of samples are below a threshold of £30,000

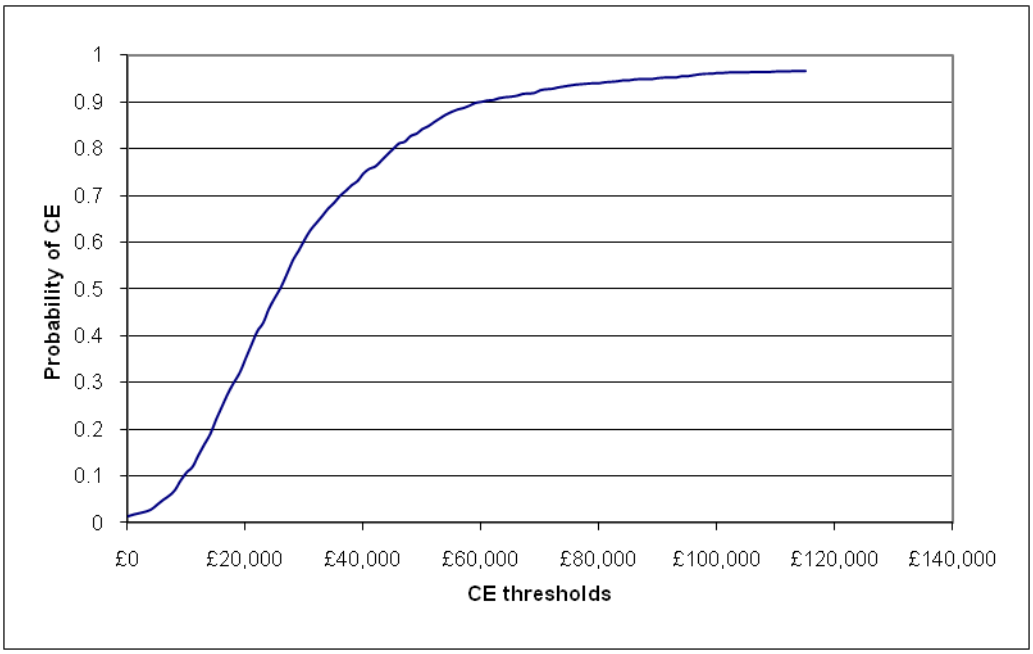




For TCZ vs. BIO

35.1% of samples are below a threshold of £20,000
60.5% of samples are below a threshold of £30,000





B9 **Priority question:** For table 87, page 236 of MS, please provide more details on the confidence interval used for withdrawal and mortality risk. Please discuss if for example the standard errors of the regression parameters have been used to assess the amount of uncertainty around the point estimate.

In the calculations for the confidence interval of the withdrawal risk it was assumed that the standard error (SE) is 30% of the mean. This is larger than the SE of the coefficient from the regression; 0.0018. In the calculations for the confidence interval of the mortality risk, the lower bound was assumed to be equal to the lowest risk in UK population for 2-18 year old patients (0.00008) [ONS 2011] . This resulted in a SE 45% of the mean. This is larger than the SE of the coefficient from the regression; 0.000015.

B10 **Priority question:** In the main analyses adjustments based on Prince 2009 are performed. Please add a sensitivity analysis using not only a variation of the Prince adjustment but use the raw Prince data as mentioned in Table 83.

In the MS Roche uses Prince et al. [2009] to adjust the efficacy of a JIA population to the subtype of systemic JIA. This is done by estimating an assumed factor of degradation based on the response of the total and the subtype populations as observed in Prince et al. [2009].

However, the results by Prince et al. [2009] as raw data are not directly comparable to the TENDER study to allow use of them as ACR response risk for the model. Prince et al. [2009] is a patient registry and not a clinical trial. The timeframe of the ACR response results do not match that of the TENDER study. Finally, Prince et al. [2009] do not include a control group to allow for an indirect comparison.

B11 **Priority question:** Both internal and external model validation are lacking. If possible please verify model outcomes with real life data (external validity), for instance using the trial data on follow up CHAQ-score. If the pivotal trials collected utility data, please validate the model against these data as well. Besides this, show the internal validity of the model for instance by performing extreme values analyses.

The TENDER clinical trial and the literature provide limited evidence for comparison with the economic model. The patient CHAQ from the TENDER trial extension at 72 weeks is provided in the table below. The number of patients in the non-response, ACR30, and ACR 50 groups at week 72 is very small and it is not appropriate to draw conclusions on the CHAQ score of those patients. With regards to the ACR 70 and ACR 90 categories, the model assumption underestimates improvement in CHAQ.

	TENDER (N)	TENDER outcomes at week 72	Model assumed values
No-response	4	0.90625	1.7442
ACR 30	3	1.291667	1.2699
ACR 50	8	1.296875	1.1351
ACR 70	20	0.63125	0.8601
ACR 90	23	0.418478	0.6692

Neither the TENDER study nor other clinical trials report utility data for comparison with the model outcomes.

In terms of internal validation, a number of extreme values were tested in the model. The table below provides a description of those tests.

Scenario	Strategy A		Strategy B		Incremental results		ICER	Justification
	Cost	QALYs	Cost	QALYs	Cost	QALYs		
The cohort in TCZ arm respond to ACR 90 (ACR 90 response= 100%)	£137,932	6.4863	£126,519	4.7161	£11,413	1.7702	£6,447	The cohort within the TCZ strategy have high response rate. The incremental treatment increases. However, the increase in health state cost exceeds that of the treatment cost resulting in lower total cost. Given the high efficacy in the TCZ arm, QALYs are significantly higher.
No response is assumed under TCZ (ACR 30 response = 0%)	£131,799	4.6679	£126,519	4.7161	£5,280	-0.0482	-£109,575	In the TCZ regimen, the cohort moves to the next treatment in the sequence after the first cycle. QALYs are very similar across the two model arms. Cost difference is driven by the initial cost of TCZ (difference of £2,400). There is a small increase in health state cost for the TCZ arm given that patients have a non-effective treatment to start with (difference £2,800).
The cohort withdraws in the first cycle of each treatment (pW=1)	£139,846	3.9030	£138,347	3.8740	£1,499	0.0290	£51,767	The cohort responders stay only for one cycle in each treatment. The difference across the two arms is driven by the proportion of ACR category response. The TCZ arm has slightly more QALYs than the comparator. There is also a minimal difference in cost driven by the no-response health state cost.

All patients die in the first cycle (pD=1)	£6,229	0.1029	£3,617	0.1029	£2,613	0.0000	Division by 0	Patients die in the first cycle at the no-response health state under the first treatment in the sequence. There is no incremental difference in QALYs. The difference in cost is driven by the treatment cost: £2,600 for TCZ and £7 for MTX. The additional health state cost (£3,600) is the same across the two model arms
The cohort in the TCZ strategy do not incur any health state cost	£89,554	5.4465	£126,519	4.7161	-£36,965	0.7304	TCZ dominates	The incremental cost difference is driven by the heath state cost of the MTX strategy
TCZ cost is 0	£77,585	5.4465	£126,519	4.7161	-£48,933	0.7304	TCZ dominates	The incremental cost difference is driven by both the treatment and heath state cost of the MTX strategy
All responder utility values are assumed 1. No response is assumed 0	£141,716	6.4341	£126,519	3.8270	£15,197	2.6071	£5,829	The results are the same with the ones presented for cost/time in a response health state
All responder utility values are assumed 1. No response is assumed 0.5	£141,716	7.8430	£126,519	6.5394	£15,197	1.3036	£11,658	The ICER increases as the difference between responder/no-responders is decreased (from previous scenario)
All utility values are assumed 1	£141,716	9.2518	£126,519	9.2518	£15,197	0.0000	Division by 0	The result is driven by the incremental cost.

Additional clarifications on cost effectiveness

B12 Please justify why in section 6.1.1 papers not in English were excluded. Please provide a reference list of those papers excluded.

The identified non-English citations are presented below. Roche had no resource to translate the studies prior to the STA.

Alekseeva,E.I.; Shakhbazian,I.E. Sandimmun-Neoral--a new quality of life for patients with severe systemic juvenile rheumatoid arthritis. 1999 Terapevticheskii arkhiv, {Ter-Arkh}, 1999, vol. 71, no. 5, p. 26-9

Brasil,Tatiana; Ferriani,Virginia; Machado,Claudia. Health related quality of life survey about children and adolescents with juvenile idiopathic arthritis. 2003 Jornal de pediatria, {J-Pediatr-Rio-J}, Jan-Feb 2003, vol. 79, no. 1, p. 63-8

Kuczynski,E.; Silva,C.A.A.; Crist□fani,L.M.; Kiss,M.H.B.; Odone,Filho,V; Assump□□o,F.B.,Jr. Quality of life evaluation in children and adolescents with chronic and/or incapacitating diseases: a Brazilian study. 2003 Anales de pediatr□a (Barcelona Spain : 2003), {An-Pediatr-Barc}, Jun 2003, vol. 58, no. 6, p. 550-5

M³ller,Godeffroy; Lehmann,H.; K³ster,R.M.; Thyen,U. Psychosocial adaptation in chronic arthritis. Behavioural characteristics of children and adolescents with juvenile idiopathic arthritis and reactive arthritis. 2006 Monatsschrift fur Kinderheilkunde, {Monatsschr-Kinderheilkd}, May 2006, vol. 154, no. 5, p. 441-447

M³ller,Godeffroy; Lehmann,H.; K³ster,R.M.; Thyen,U. Quality of life and psychosocial adaptation in children and adolescents with juvenile idiopathic arthritis and reactive arthritis. 2005 Zeitschrift f³r Rheumatologie, {Z-Rheumatol}, Apr 2005, vol. 64, no. 3, p. 177-87

Martini,F. Quality of life in pediatric rheumatology. 2000 Archives de p□diatrie : organe officiel de la Soci□te fran□aise de p□diatrie, {Arch-Pediatr}, May 2000, vol. 7 Suppl 2, p. 233s-234s

Petersen,C.; Nordmeyer,S.; M³ller,Godeffroy; Foeldvari,I.; K³ster,R.M.; Bullinger,M. Health-related quality of life in children and adolescents with juvenile idiopathic arthritis: which role do age, sex and medical parameters play? 2008 Klinische P□diatrie, {Klin-Padiatr}, Jul-Aug 2008 (epub: 12 Feb 2008), vol. 220, no. 4, p. 259-65

Reicher,E.; Reipert,D.; Tempska,K.; Kowalczewska,J. Socioeconomic evaluation of results of treatment and rehabilitation of children and adolescents treated for rheumatoid arthritis at the Pediatric Center of the Institute of Rheumatology at Konstancin as compared with the results of treatment of chronically ill children in Social Welfare Institutions. 1974 Reumatologia, {Reumatologia}, 1974, vol. 12, no. 3, p. 199-206

Salugina,S.O. Retrospective analysis of the course and outcomes of juvenile arthritis in adult patients. 2010 Terapevticheskii Arkhiv, {Ter-Arkh}, 2010, vol. 82, no. 5, p. 22-29

Van,Herwaarden; Ten,Berge; Prakken,A.B.J.; Sinnema,G.; Kuis,W.; van,der The effects of polyarticular juvenile chronic arthritis in everyday life. 1997 Nederlands Tijdschrift voor Geneeskunde, {NED-TIJDSCHR-GENEESKD}, 1997, vol. 141, no. 4, p. 192-196

B13 Please confirm that in table 74, page 213 MS, ACR-response refers to the 12 week follow up measurement of patients.

Please clarify the table number. Table 74 does not contain any ACR response data.

B14 Please confirm that the probabilities mentioned in tables 80 and 82, pages 229 and 234 MS, refer to ACR response as determined after 12 weeks of treatment.

Table 80 summarises the evidence of withdrawal risk for biologic treatments. Table 81 presents the ACR response data after 12 weeks of treatment for TCZ and MTX arms. Table 82 presents the assumed probabilities for the TCZ and MTX arms based on table 81 and after the transformation described in section 6.3.2 in the MS.

B15 Please explain what the word 'sampling' refers to in table 87, page 236 MS, related to utility values.

The value used in the PSA is not independently sampled by another distribution but linked to the probabilistic value of the ACR-No-Response utility parameter.

B16 The costs associated with the heath states show a very large difference between patients without response and patients with response. While the difference between ACR90 and ACR30 is only 171.44, the difference between ACR30 and no response is 3094.91. Please justify these differences, considering that the only source of information was expert opinion.

Roche has identified that health state cost is a driver of the model cost-effectiveness results (section 6.10.3). Due to lack of evidence on resource use clinical expert opinion was sought in order to identify the appropriate values to populate the model.

Of the resource use items, the cost of inpatient stay, and in particular the length of hospitalisation, has the biggest impact on the results. All clinical experts suggested that if the analysis considers a week of hospitalisation for the average patient who responds to treatment then assuming a month for non-responders is an underestimate. Given the severe symptoms of the patient's condition, namely fever and skin rash, and also considering that all patients are of young age, it seems plausible to suggest that extensive hospitalisation would occur while patients experience disease flare. It is also expected that non-responders would experience a number of disease flares in a given year. Clinical experts suggested that for non-responders the length of hospitalisation could far exceed three weeks and that the average patient could stay in hospital for as much as three months in a year. The analysis has taken a conservative approach on this estimate and considers the lowest value suggested by clinical experts (3-4 weeks a year).

B17 Related to the previous question please justify the remark in table 97, index 21, that the current assumptions about cost of inpatient stay are conservative.

Please see response to B16; the approach is conservative against what clinical experts consider a more plausible value.

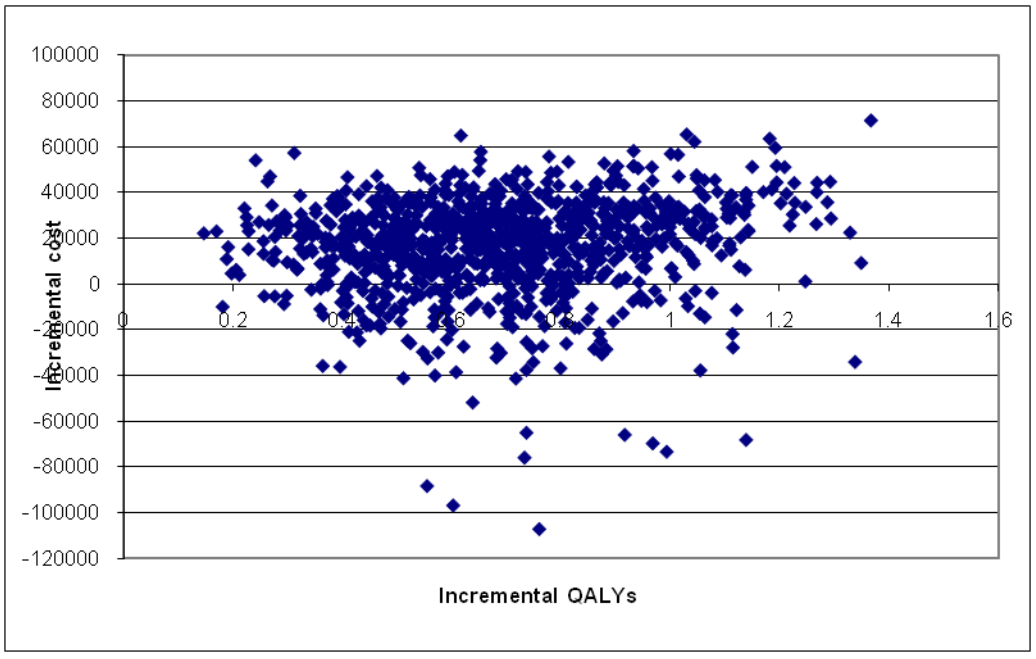
B18 Please provide a new PSA with larger margins around the duration of hospitalisation and the percentage of patients requiring hospitalisation.

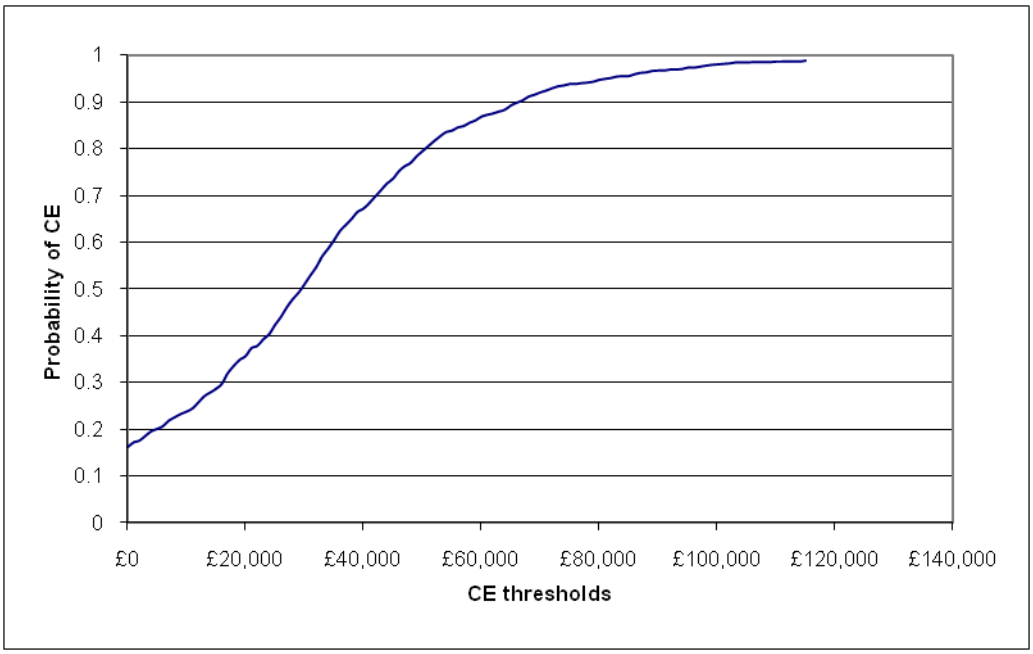
The economic model was modified and a margin of 80% of the mean was assumed as standard error. The PSA results presented below include the changes to the Dirichlet distribution as requested from question B8.

Parameter	Alpha (Gamma dist)	Beta (Gamma dist)
Hospitalisation of responder	1.5625	4.8
Hospitalisation of non-responder	1.5625	15.68

TCZ vs. MTX:

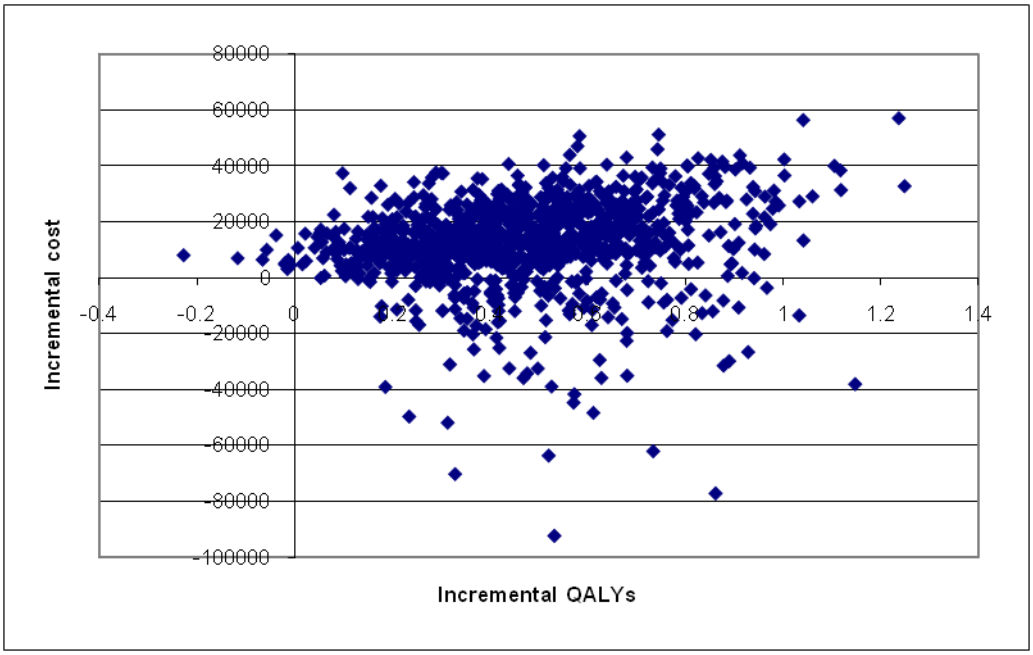
35.7% of samples are below a threshold of £20,000
51.1% of samples are below a threshold of £30,000

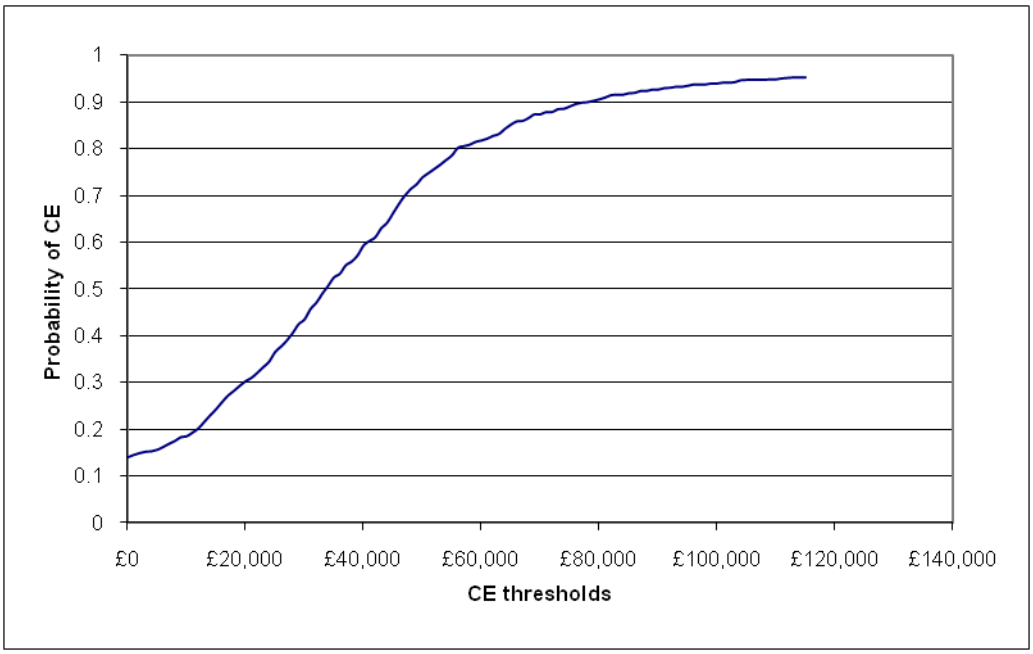




For TCZ vs. BIO

30.2% of samples are below a threshold of £20,000
43.4% of samples are below a threshold of £30,000





Literature searches

B19 For section 6.1.1 please clarify whether the required database, EconLIT, was searched.

EconLIT was not searched for the original submission. Roche performed this search and details can be found in the Appendix of the literature searches.

B20 For section 9.10.2 please provide the date HEED was searched.

The strategy for the search performed on the 3rd of December 2010 was not recorded. Roche repeated the review of HEED. Please refer to the accompanying appendix on literature searches for the details of the new search.

B21 According to page 202 of the MS, HEED was searched to identify cost-effectiveness evaluations. Please provide the full search strategy, date searched and date span for the HEED search.

Please see above (question B20).

B22 Please provide the full search strategy for the NHS EED search.

The strategy for the search performed on the 3rd of December 2010 was not recorded. Roche repeated the review of NHS EED. Please refer to the accompanying appendix on literature searches for the details of the new search.

B23 Section 6.4.5 states that the cost-effectiveness searches details in 6.1.1 were also used to retrieve HRQL references. Please explain why the search strategies

described in 6.1.1 and 9.10 do not include terms such as ‘HRQOL’ or HRQL instruments such as EuroQOL, EQ5D or SF36.

The search strategies for cost-effectiveness studies are based on the search filter for economic evaluation developed by CRD. The search includes the terms quality of life, health status, quality adjusted life years, and well being. Roche considered adequate the inclusion of the above terms only, without adding the terms mentioned here.

The updated searches presented for this response (EconLIT and NHS EED) included the additional terms but did not yield to any additional missing studies.

REFERENCE:

Mahesh K. B. Parmar, Valter Torri and Lesley Stewart. 1998. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Statistics in Medicine 17:2815-2834.

Enclosures

Copy of this response with AIC/CIC removed

Appendix of literature reviews

[redacted]

Sent out in the post

CD with requested ‘additional data’